DOCUMENT RESUME

ED 156 677 TM 007 065

AUTHOR Chan, Tak-Cheung
TITLE Determining Valid Criteria for Teacher Evaluation.
PUB DATE [77]
NOTE 18p.
JOURNAL CIT Georgia Association of Middle School Principals' Journal; v1 n1 Spr 1977

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Effective Teaching; *Evaluation Criteria; Predictor Variables; *Reliability; *Research Needs; Teacher Behavior; *Teacher Characteristics; *Teacher Evaluation; Teacher Qualifications; Teaching Skills; *Validity

ABSTRACT

Teaching may seem too complex and too involved in various dimensions of affective and cognitive aspects of learning for a good teacher evaluation program to be devised. A hierarchy of criteria is proposed, based upon analysis of observable versus unobservable characteristics, measureable and non-measureable characteristics, and attributes and variables. Once reliability is established, then the conceptually more difficult task of determining validity can begin. The ultimate goal of validity can only be based upon reliability. The research task of developing reliable methods of evaluating teachers and determining the validity of the criteria on which these methods rest lies before us. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED156677

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

DETERMINING VALID CRITERIA FOR TEACHER EVALUATION

Tak-Cheung Chan

Research Assistant

Georgia Finance Study

College of Education

University of Georgia

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Carvin L. Brown
GAMSPJ Editor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

One of the issues in teacher evaluation that needs to be resolved, possibly before all others, is that of determining the criteria for measuring teacher effectiveness. The criteria depend on a high degree of validity to be useful to a teacher evaluation program. This article attempts to develop a foundation prerequisite to the examination of the validity of evaluation criteria and to indicate direction for further research on teacher evaluation.

Method of Review

A review of the research literature on teacher evaluation criteria was made with reference to the Research in Education (ERIC) - January, 1972 to December, 1974 and the Resources in Education (ERIC) - January, 1975 to November, 1976. Simon and Boyer's Mirrors for Behavior: An Anthology of Classroom Observation (1967) and Gage's Handbook of Research on Teaching (1963) were widely used as additional references. For an overall

TM007 065

survey of the educators' efforts on teacher evaluation, Barr's Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness (1961) and Biddle and Ellena's Contemporary Research on Teacher Effectiveness (1964) were frequently referred to. In addition, Dubin's Theory Building (1969) was used as a basis for structuring the theoretical background of this article.

Findings

In the Preface to Contemporary Research on Teacher Effectiveness, Biddle and Ellena (1964) point out that:

> ... with all this research activity, results have been modest and often contradictory. Few, if any, facts are now deemed established about teacher effectiveness; and many former findings have been repudiated. It is not an exaggeration to say that we do not today know how to select, train for, encourage or evaluate teacher effectiveness (Preface, p. vi).

In the Evaluation of Teacher Effectiveness, McDaniel (1972), after reviewing Biddle and Ellena's remark, says:

> Considering the present state of art, the temptation is large for becoming a prophet of doom crying out against further efforts to undertake so fruitless a task as teacher

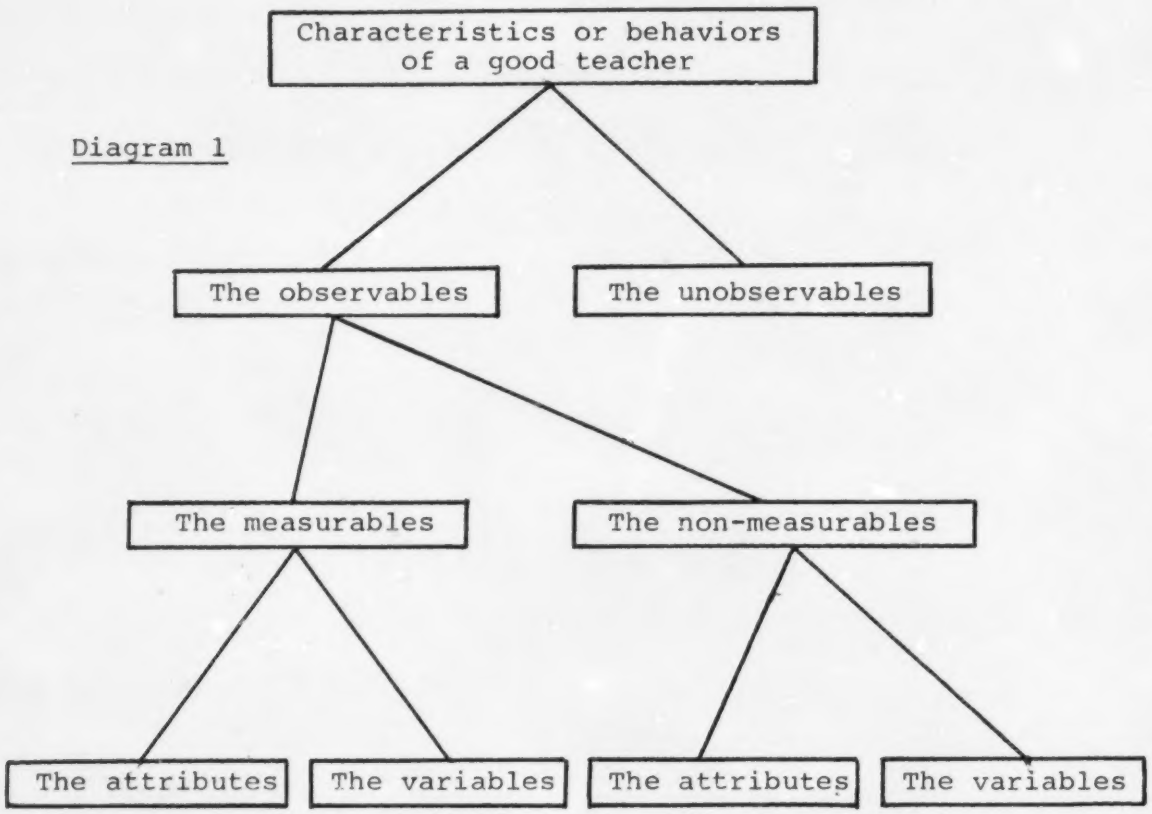
evaluation (p. 1).

Even though the process of teacher evaluation as described by these educators is very complicated, it is certainly not a dead-end process. In fact, the effectiveness of the evaluation program depends to a great extent on the validity of the criteria used in the evaluation of teachers. Therefore, program developers must decide what the teachers are to be evaluated upon; in other words, what the evaluation criteria are to be. In addition, they must be sure that the criteria they choose are valid. Only valid criteria can result in meaningful teacher evaluation.

The "Good" Teacher

Educators have identified the characteristics or behaviors of good teachers (e.g., A.S. Barr, 1958 and D. Hamacheck, 1968). When these characteristics or behaviors are perceived from an evaluation point of view, they can be divided into observables and unobservables with the observables further subdivided into attributes and variables, as well as measurables and non-measurables (see: Diagram 1).

Diagram 1



The best evaluation program that can be devised is limited to observable characteristics or behaviors because no matter how carefully we observe, there are still good teacher characteristics or behaviors we cannot observe. We can identify a list of characteristics or behaviors of good teachers, but we cannot necessarily expect a person who possesses these characteristics or behaviors to be a good teacher. As Dubin (1969) says:

> We simply are not capable of seeing things whole. Nor is man capable of retaining and recording complex phenomena coming within the range of his sensory fields. It is necessary to acknowledge that man, who builds theories to model his world of observation, has genuine limits on his capacities to grasp complex observations (p. 30).

The measurables are those characteristics or behaviors that can be measured by special devises in relation to criteria. The non-measurables are those that cannot be measured by any device or in any other form except as subjective opinion. Many educators have been interested in the measurability of teacher characteristics. Wallen and Travers (1963) point out that "even in those situations defined strictly in terms of observables, such observables are themselves

abstractions..." (p. 448) Redfern (1963) refers to a teacher's personality traits, habits, manner, judgment, emotional stability, initiative, tact and courtesy as being hard to evaluate. Gibson and Hunt (1965) also state that educational outcomes are often hard to measure. Stoops, Rafferty and Johnson (1975) talk about pupil subject achievement being measurable while social, moral, and emotional progress are relatively intangible. On the whole, good teacher characteristics or behaviors contain abstractions which are non-measurable.

In defining an attribute and a variable, Dubin (1969) states:

> An attribute is a property of a thing distinguished by the quality of being present. The thing always has this quality if the attribute is a property of the thing. All things having a given attribute property constitute a set of identities on that attribute property. All other things are in a set identified by the lack of the given attribute property. A variable is a property of a thing that may be present in degree. There may be some of the property present or a lot of it... What is significant when we employ a variable unit in a theory

is that our attention becomes focused upon the amount or degree to which this property is present in the thing (p. 35).

In other words, by attributes, we mean the generally common characteristics or behaviors that can be identified as existing in good teachers. By variables we mean the quantity of characteristics or behaviors that are required of individual teachers. In the sense of essentiality, attributes constitute an inevitable part of the characteristics or behaviors of good teachers, while variables are present in different degrees of importance.

Establishing Reliability

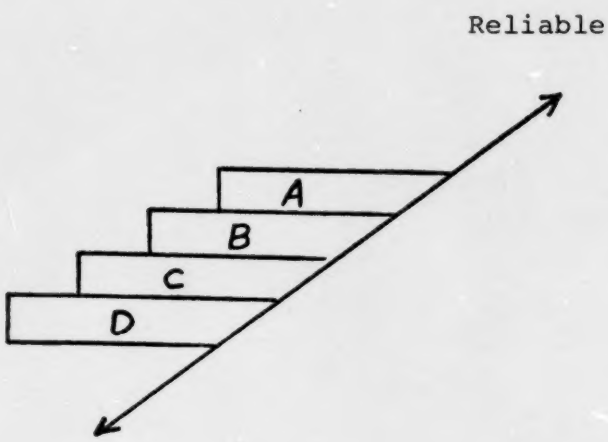
Establishing teacher evaluation criteria based upon good teacher characteristics or behaviors is a common practice of the school administrators. However, the question of the validity of these characteristics or behaviors remains unanswered.

To examine the validity of teacher evaluation criteria, an attempt is made to design a hierarchy of reliability as follows:

If the measurable attributes are represented by A, the measurable variables by B, the non-measurable attributes by C, and the non-measurable variables by D; and if it is assumed that as evaluation criteria,

the attributes are more reliably-observed than the variables and the measurables are more reliably-observed than the non-measurables, then A, B, C and D can be arranged in a simple hierarchy of reliability (see: Diagram 2) for examining teacher evaluation criteria with D, C, B and A in an ascending order of reliability.

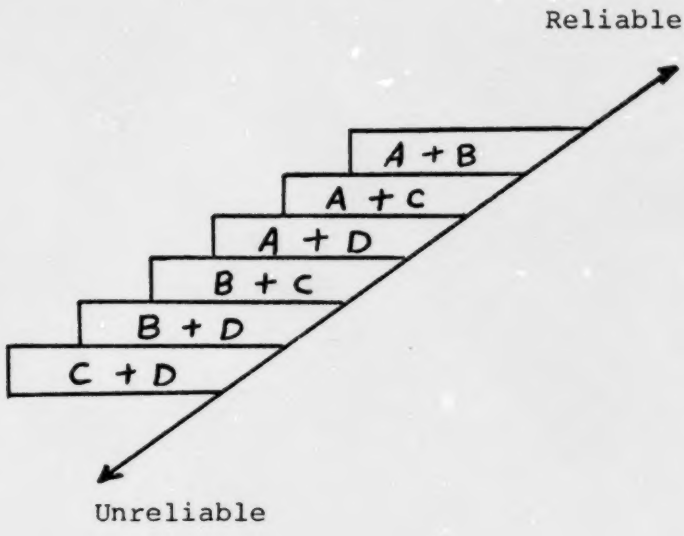
Diagram 2



In evaluation programs, the criteria are often based on more than one type of teacher characteristics or behaviors. Therefore, there are possibilities of two, or three or even four types of teacher characteristics or behaviors involved. In considering the involvement of only two types of teacher characteristics or behaviors, a more complicated hierarchy of reliability can be constructed through pairings (see:

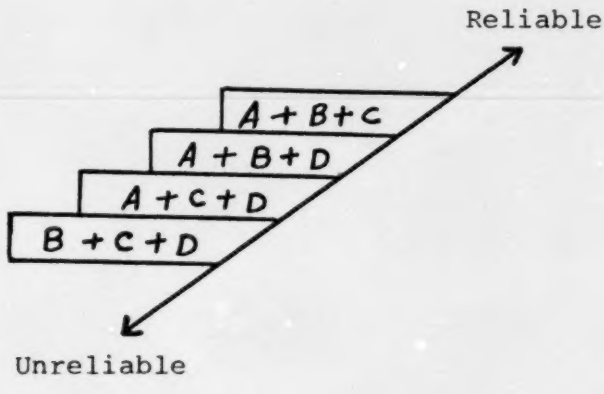
Diagram 3) with (C + D), (B + D), (B + C), (A + D), (A + C) and (A + B) in ascending order of reliability.

Diagram 3



In the same way, in the involvement of three types of teacher characteristics or behaviors, another more complicated hierarchy of reliability can be constructed through triads (see: Diagram 4) with (B + C + D), (A + C + D), (A + B + D) and (A + B + C) in ascending order of reliability.

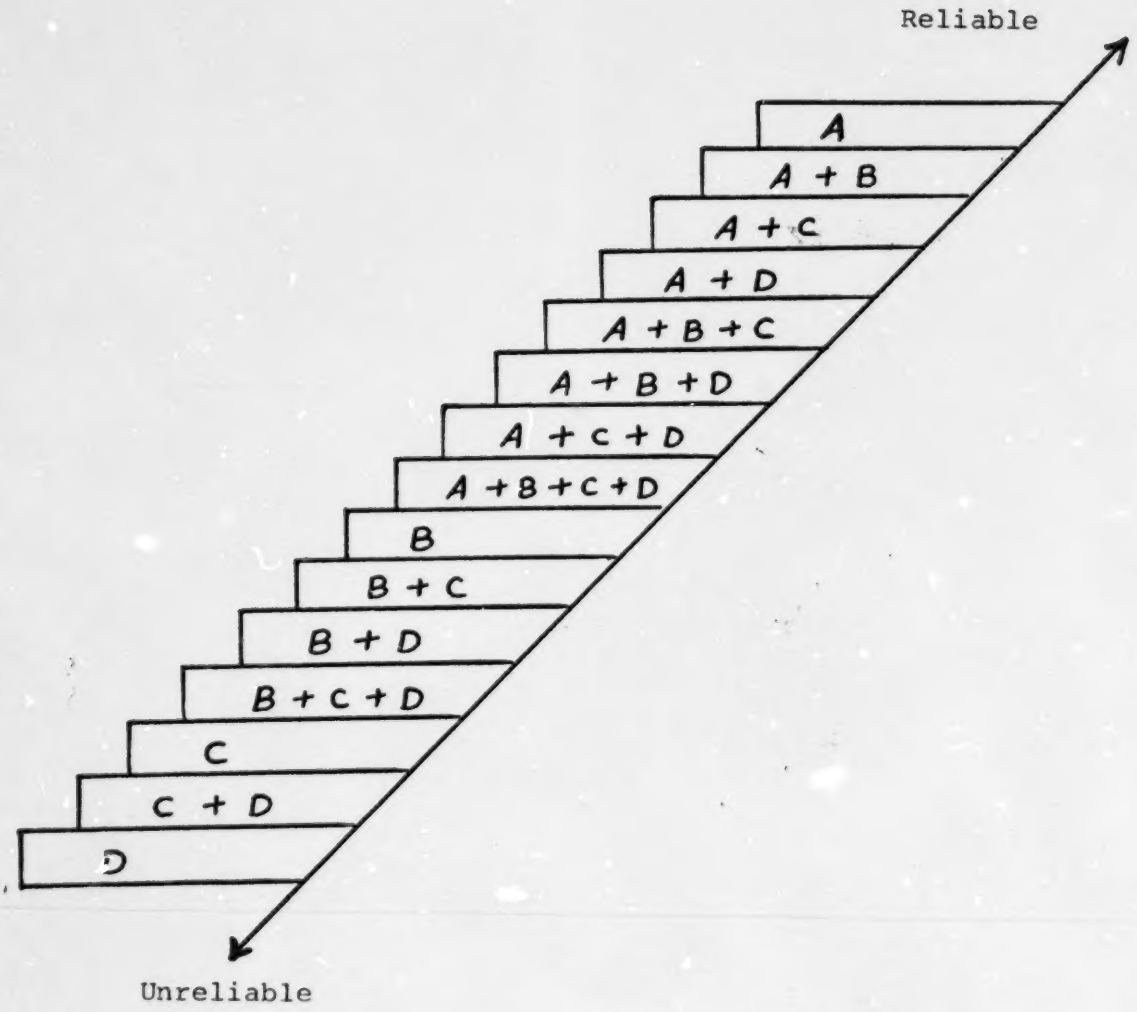
Diagram 4



Moreover, the utilization of four types of teacher characteristics or behaviors can be expressed as (A + B + C + D).

By using the same previous assumptions, considering the possibility of A + B + C + D, and further assuming that the components in the pairs, triads and the quartet are proportionally equal in quantity (e.g., A + B means A and B are proportionally equal in quantity), then diagrams 2, 3 and 4 can be logically combined to derive a more complete fifteen level hierarchy of reliability (see: Diagram 5).

Diagram 5


Reliable
A
A + B
A + C
A + D
A + B + C
A + B + D
A + C + D
A + B + C + D
B
B + C
B + D
B + C + D
C
C + D
D
Unreliable

The simple hierarchy in Diagram 2 can be used to examine the reliability of individual teacher evaluation criterion. The complicated hierarchies in Diagram 3 and 4 can be employed as a means to compare the degree of reliability among evaluation programs involving "pairs" and "triads" characteristics or behaviors. As to the fifteen level hierarchy in Diagram 5, it is devised to examine the reliability of any teacher evaluation program.

Needed Research

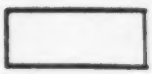
More research is needed to clear up many unsolved problems relative to the reliable teacher evaluation methods. Some of the pressing problems are given below:

In constructing the hierarchy of reliability (Diagram 5), it is assumed that the attributes and the variables can be identified. But, in actual practice, to distinguish attributes from variables involves much personal judgment. However, since "an attribute is a property of a thing distinguished by the quality of being present and all other things are in a set identified by the lack of the given property," the attributes and variables can be described. In Diagram 6, the three circles stand for three different evalu-

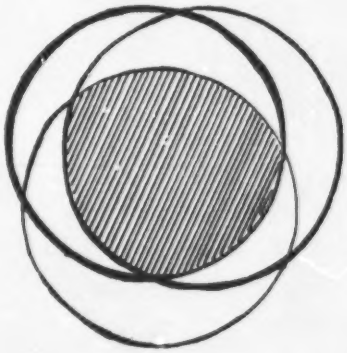
ation programs, the shaded area represents the attribute criteria, and the unshaded areas, the variable criteria.

Diagram 6

attributes



variables



In other words, criteria that are commonly used in many evaluation programs are attribute criteria and criteria that are not commonly used in evaluation programs are variable criteria. As long as it is difficult to determine the degree of commonness, the problem of attributes and variables leave much to be researched.

Through years of research, educators keep admitting that there are teacher characteristics or behaviors that cannot be measured. But intellectual advancements have been able to make many things possible. Concerned educators have never ceased their efforts in attempting to experiment with new strategies of evaluation. Therefore, evaluation program planners should be aware of the fact that regardless of the new developments making many

teacher characteristics or behaviors measurable, there is still much needed research in the area of the non-measurables.

It is assumed in Diagram 5 that the component criteria in the pairs, the triads and the quartet (e.g., A + B, A + B + C, and A + B + C + D) are proportionally equal in quantity. But in actual practice, the component criteria are frequently not quantitatively equal in proportion. Appropriate adjustments are required. There is a need for studies of the degree of appropriate adjustments.

Conclusion

Teaching may seem too complex and too involved in various dimensions of affective and cognitive aspects of learning to devise an universal good teacher evaluation program. However, teaching can be evaluated in a meaningful sense if, and only if, we can establish reliability in our methods. Although the hierarchy of reliability (Diagram 5) is not a technically satisfactory device, it is expected that the idea can suggest other alternative directions to a more practical strategy of teacher evaluation.

Once reliability is established then the conceptually more difficult task of determining validity can begin. The ultimate goal of validity can only be based

upon reliability. The research task of developing reliable methods of evaluating teachers and determining the validity of the criteria on which these methods rest lies before us.

REFERENCES

Barr, A.S. Characteristics of successful teachers. Phi Delta Kappan, 1958, 39 (6) 282-283.

Barr, A.S. Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations. Madison, Wisconsin: Dembar Publications Inc., 1961.

Biddle, B.J., and Ellena, W.J. Contemporary Research on Teacher Effectiveness. New York: Holt, Rinehart and Winston, 1964.

Brighton, S., and Rose, G. Increasing your Accuracy in Teacher Evaluation. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1965.

Dubin, R. Theory Building. New York: The Free Press, MacMillan Co., 1969.

Gibson, R.O., and Hunt, H.C. The School Personnel Administrator. Boston: Houghton Mifflin Co., 1965.

Hamachek, D. Characteristics of good teachers and implications for teacher education. Phi Delta Kappan, 1969, 50 (6) 341-344.

McDaniel, E. Evaluation of Teacher Effectiveness. Washington, D.C.: Office of Education, U.S. HEW Department, 1972.

Redfern, G.B. How to Appraise Teaching Performance. Columbus, Ohio: School Management Institute, Inc., 1963.

Simon, A., and Boyer, E.G. Mirrors for Behavior: An Anthology of Classroom Observations Instruments. Philadelphia: Research for Better Schools, Inc., 1967.

Smith, B.O. Teaching: Conditions of its evaluation. In The Evaluation of Teaching. Washington, D.C.: Pi Lambda Theta, 1967.

Stoops, E., Rafferty, M. and Johnson, R.E. Handbook of Educational Administration. Boston, Massachusetts: Allyn and Bacon, Inc., 1975.

Wallen, N.E., and Travers, R.W.M. Analysis and investigation of teaching methods. In Handbook of Research on Teaching., a project of the American Educational Research Association. Chicago: Rand McNally, 1963.